



Diseño y validación de una rúbrica para evaluar la adecuación de artículos científicos al estilo de la American Psychological Association

Gladys Merma Molina¹, Hilda Peña Alfaro², Silvia Rosa Peña Alfaro González³

¹ Departamento de Didáctica General y Didácticas Específicas, Facultad de Educación, Universidad de Alicante, España {gladys.merma@ua.es}

² New World Spanish-Comunicart, Estados Unidos {hilda@nwspanish.com}

³ Comunicart, Consultoría en Lingüística Aplicada, México {comunicart@live.com}

Recibido el 31 Octubre 2016; revisado el 14 Noviembre 2016; aceptado el 15 Noviembre 2015; publicado el 15 Enero 2017

DOI: 10.7821/naer.2017.1.220



RESUMEN

En este estudio se expone el proceso de diseño y validación de una rúbrica para evaluar la adaptación de los artículos científicos a las normas de la *American Psychological Association* (APA). La rúbrica evalúa ciertos aspectos del estilo APA que les permitan a los autores, editores y evaluadores valorar si el artículo científico es coherente con dichas normas. En concreto, la rúbrica se concentra en los *Aspectos Generales* del artículo y en el *Sistema de citación*. Para ello se analizaron 10 artículos publicados entre los años 2012-2016 en revistas incluidas en la *Journal Citation Report*, empleando la *técnica de juicio de expertos*. Después de haber realizado 5 pruebas piloto, los resultados mostraron la validez y confiabilidad del instrumento. Además, el proceso puso en evidencia las posibilidades de la rúbrica para contribuir a la uniformidad de criterios divergentes y para emplearse como herramienta didáctica en diferentes escenarios.

PALABRAS CLAVE: INVESTIGACIÓN, CIENCIA, ESTANDARIZACIÓN, MEDICIÓN

1 INTRODUCCIÓN

El *Manual de Publicaciones de la American Psychological Association* (en lo sucesivo el *Manual*) se publicó por primera vez en el año 1929 como un documento de siete páginas (APA, 2010) que ofrecía un conjunto de reglas para definir cómo debía realizarse la presentación de información científica. A partir de entonces ha venido actualizándose hasta llegar a la última edición, que es la sexta en inglés y la tercera en español. Aunque tradicionalmente se le identifica por su sistema de citación, el *Manual* es una valiosa herramienta y “una fuente de autoridad en todos los sentidos de la escritura académica” (APA, 2010, p. 3) ya que ofrece una serie de normas y orientaciones para lograr la máxima precisión en la escritura académica y para identificar las mejores prácticas en el manejo ético y responsable del manuscrito a lo largo del proceso de publicación. Si bien el *Manual* ofrece normas específicas para los diversos aspectos de la escritura académica, según la visión de las expertas, participantes en esta

investigación, existen serias deficiencias en la aplicación del estilo APA que entorpecen el proceso de revisión de los artículos científicos.

La normalización en las publicaciones periódicas facilita la comunicación y difusión científica, por lo que es importante unificar los formatos y procedimientos de edición de los artículos de investigación con el fin de incrementar la transferencia, el intercambio, el acceso y la disponibilidad de la información. El proceso de selección y evaluación de los trabajos de investigación mediante la *revisión por pares* o *arbitraje*, utilizado por las revistas científicas o publicaciones periódicas, es un recurso para asegurar que los artículos cumplan con ciertos estándares en cuanto a contenidos y estilo. Este proceso representa uno de los pilares que distingue a la literatura científica de calidad (Giménez, 2014).

Medir la calidad de un trabajo de investigación es un proceso complejo, pero existen variables comunes que permiten realizar una valoración homogénea. Floyd et al. (2011) señalan diferentes razones por las que los manuscritos que se presentan a las revistas para su evaluación son rechazados y sostienen que la selección de dichos trabajos por los editores se centra tanto en cuestiones generales como en otras más específicas. Entre las primeras figuran la temática de la investigación, la metodología, la escritura y el formato del manuscrito, y en la revisión especializada, que es la evaluación de los revisores expertos, se valora especialmente la calidad del contenido de los manuscritos, su vinculación con la investigación previa, sus contribuciones a la ciencia y si estas son científicamente sólidas. No obstante, tanto los editores como los evaluadores pueden reaccionar negativamente ante un manuscrito sólido que muestre signos de prisa o de falta de cuidado en la preparación (Nickerson, 2005). Asimismo, las deficiencias en la escritura y en el formato son razones comunes para que los manuscritos sean devueltos a los autores antes de haber pasado por una revisión completa. Por ello, coincidimos con Camacho, Rojas y Rojas (2014), quienes sostienen que el autor debe realizar una autoevaluación, tanto en contenido como en forma, antes de enviar el artículo a una revista; esto significa que es importante asegurarse de que las ideas se hayan planteado de modo claro y coherente, procurando un estilo depurado accesible, interesante y comprensible para el lector.

Determinar qué manuscritos son aceptables para su publicación de forma objetiva requiere de un buen método, tanto para los

*Por correo postal dirigirse a:

University of Alicante, Carretera de San Vicente del Raspeig, s/n
03690, Alicante, Spain

autores como para los equipos editoriales de las revistas, y, además, evitaría las suspicacias de estos últimos respecto a que el proceso de evaluación no haya sido del todo objetivo ni transparente. En ese sentido, es necesario, especialmente en el ámbito de las ciencias sociales, contar con un instrumento fiable y confiable, basado en una escala cuantitativa como es la rúbrica.

Es importante destacar que el término *rúbrica* es usado de maneras muy diferentes en el ámbito académico. En esta línea, Hafner y Hafner (2003) argumentan que el concepto es confuso, pero que en el entorno académico se usa para describir un instrumento de evaluación que permite valorar los niveles de desempeño de una tarea en particular y los resultados basados en contextos específicos.

La rúbrica analítica propuesta por Moskal (2000) y Moskal y Leydens (2000) para el desarrollo de la evaluación en el ámbito educativo garantiza la promoción de prácticas de evaluación más justas (Picón, 2013). Andrade (2005) y Peat (2006) destacan la importancia del uso de la rúbrica porque brinda criterios definidos y explícitos que conducen a una mayor objetividad en la evaluación y permite valorar varias dimensiones. En el ámbito de la estandarización de la investigación, la rúbrica ha sido escasamente utilizada; solo en los últimos años se han comenzado a ver los esfuerzos por crear matrices de valoración para el propósito específico de los informes de investigación especialmente de carácter empírico (Greenberg, 2012).

Las exiguas experiencias de la rúbrica en la valoración de los artículos de investigación evidencian la necesidad de elaborar un instrumento de medición que demuestre su fiabilidad y validez: la fiabilidad, en el sentido de la coherencia en los resultados, se realiza con respecto al cálculo de los evaluadores a través de una descripción y análisis cuidadoso de criterios específicos uniformes mediante el cual se logra que el papel de las opiniones subjetivas se reduzca al mínimo (Newell, Dahm, & Newell, 2002). La medición fiable de un instrumento, tal como señalan Hernández, Fernández y Baptista (2003), se refiere al grado en que su aplicación repetida al mismo sujeto u objeto produce iguales resultados. Por su parte, la validez o la repetibilidad, como arguyen Moskal y Leydens (2000), Stellmack et al. (2009) y Thaler, Kazemi y Huscher (2009), es el grado en que un instrumento realmente mide variables subyacentes de interés. Finalmente, en la literatura asociada al diseño de la rúbrica aparece el término *precisión* (Brown, Glasswell, & Harland, 2004), el cual se refiere a la relación psicométrica de calidad entre los indicadores de fiabilidad y validez, lo que normalmente se expresa como un porcentaje exacto de correlación entre ambos factores (Cho, Schunn, & Wilson, 2006; Jonsson & Svingby, 2007).

En base a estos antecedentes, el objetivo de este estudio es diseñar y validar un instrumento para evaluar la adecuación de los artículos científicos a las normas APA.

2 MÉTODO

Para el diseño y validación de la rúbrica se utilizó la *técnica de expertos*, que es la base del Método Delphi (Steurer, 2011). Esta técnica, que favorece la creación de directrices (Roddy et al., 2006), se ha empleado especialmente en la medicina, en las políticas públicas, en la evaluación educativa y en las ciencias sociales (De Villiers, M. R., De Villiers, P. J., & Kent, 2005). Se sustenta en la selección de expertos cualificados que analizan una temática, un instrumento o una cuestión de investigación. La técnica de expertos incluye comentarios, informes y juicios que son anónimos, donde cada especialista aporta su *opinión de*

experto. Estos juicios se recopilan en un proceso sistemático que se centra en puntos de vista personales, y que luego se consensuan a nivel de grupo (Steurer, 2011). Entre las ventajas de la técnica destacan la posibilidad de que los expertos puedan ser de diferentes partes del mundo, que puedan participar vía electrónica y que no haya una influencia mutua entre ellos (Burt et al., 2009). Teniendo en cuenta estos principios, en este estudio participaron tres especialistas en las normas de la APA con amplia experiencia en los procesos de evaluación y revisión de artículos científicos y en la edición de revistas, tanto en inglés como en español.

Se analizaron 10 artículos publicados entre los años 2012-2016 en distintas revistas de habla hispana del área de Ciencias Sociales incluidas en el Institute for Scientific Information (ISI), actualmente Thomson Reuters o Journal Citation Reports, en cuyas "instrucciones para los autores" se recomienda el estilo APA como guía de referencia para su elaboración. Los factores de impacto de Thomson Reuters proporcionan un índice del efecto de los artículos que aparecen en una revista durante un periodo de tiempo determinado, normalmente dos años anteriores al cálculo; dichos indicadores adquieren sentido si se comparan con los indicadores de otras revistas de la misma temática.

3 PROCESO DE DISEÑO DE LA RÚBRICA

Primera etapa

Inicialmente se planteó como objetivo diseñar y validar una rúbrica para evaluar artículos científicos conforme al estilo APA. Era claro, sin embargo, que dicho instrumento no podría abarcar todos los aspectos de la normativa propuestos por el *Manual* y que había que efectuar una selección de los factores más importantes del estilo APA. En principio, se consideró que el sistema de citación era un elemento ineludible por ser la columna vertebral del estilo APA. El sistema de citación se refiere al cuerpo de pautas especificadas en el capítulo 6 del *Manual*, bajo el título de *Citación de fuentes* (APA, 2010, p. 169-192). Este conjunto de reglas incluye la citación directa de las fuentes y la paráfrasis, la citación de referencias en el texto y en la lista de referencias.

Tabla 1. Categorías de los Aspectos generales y del Sistema de citación

Aspectos generales	Sistema de citación
Título	Citación directa y paráfrasis
Secciones	Referencias en el texto
Nombre del autor afiliación institucional	Lista de referencias
Resumen	
Introducción	
Método	
Resultados	
Comentarios (Discusión)	
Estructura del artículo	
Lenguaje empleado libre de discriminaciones	

Además del sistema de citación, se determinó que era necesario añadir otros aspectos que resultan cruciales en la escritura científica como las secciones del artículo, su estructura y el empleo de un lenguaje libre de discriminaciones. Así, se consensó que la rúbrica consideraría dos apartados: *Aspectos generales* y *Sistema de citación* con sus correspondientes categorías. La Tabla 1 ilustra la manera como se dispusieron cada una de las categorías a evaluar.

El siguiente reto fue identificar los parámetros para evaluar cada una de las categorías. Con respecto a los *Aspectos generales*, se acordó que cada uno de ellos debía ceñirse de manera estricta a los lineamientos del *Manual*. De acuerdo con esto, se estipularon los siguientes parámetros:

Tabla 2. Parámetros para evaluar los *Aspectos generales*

1.	El título sintetiza la idea principal del escrito de una manera sencilla y con un estilo adecuado
2.	El artículo contiene las secciones propuestas por el estilo APA: resumen, introducción, método, resultados, comentarios (discusión).
3.	El nombre de autor y la afiliación institucional se ciñe a los criterios indicados por el estilo APA.
4.	El resumen es preciso, no evaluativo, coherente, legible y conciso.
5.	La introducción plantea el problema y describe la estrategia de la investigación.
6.	La sección de método incluye una descripción completa de los métodos empleados.
7.	La sección de resultados presenta los datos relevantes que sustentarán las conclusiones.
8.	La sección de comentarios (discusión) examina, interpreta y califica los resultados, enfatizando sus consecuencias teóricas o prácticas.
9.	El artículo tiene una estructura, caracterizada por la utilización de diferentes niveles de encabezados para indicar la jerarquización de la información.
10.	El lenguaje empleado está libre de discriminaciones a grupos o a personas.

En relación con las Secciones (Parámetro 2), el *Manual* propone incluir el resumen, la introducción, el método, los resultados, los comentarios (discusión) y las referencias. Efectivamente, es inobjetable la conveniencia de contar con dichos apartados ya que sientan las bases para una distribución organizada de la información en el artículo, que resulta además una guía muy útil para el lector. No obstante, el *Manual* no estipula que los artículos científicos deban contar estrictamente con dichos apartados. De hecho, afirma: “El *Manual de publicaciones* presenta requerimientos de estilo explícitos, pero reconoce que las alternativas a veces son necesarias. Los autores deben equilibrar las reglas del *Manual de publicaciones* con el buen juicio” (APA, 2010, p. 5).

Así pues, se determinó que, sin importar qué nombre se les diera a las secciones, lo crucial era que el artículo presentara los apartados necesarios para dar cuenta de los resultados de la investigación y que la información estuviera debidamente organizada para facilitar la comprensión y la comunicación. De aquí se concluyó que el revisor tendría que aplicar su “buen juicio” para evaluar este aspecto. Una vez delimitadas las categorías y los parámetros, se determinó fijar una escala de 0 al 5 para evaluar cada una de las categorías.

Segunda etapa

Con esta base se realizó la Prueba piloto 1, que consistió en que las 3 expertas evaluaran de manera independiente los artículos A, B y C en lo concerniente a la categoría de *Aspectos generales*. Considerando que dicha categoría contaba con 10 parámetros, se determinó una escala de evaluación de 0/10 puntos. La prueba dio como resultado que había un 65% de acuerdos en la puntuación alcanzada contra un 35% de desacuerdos. Si bien se contaba ya con criterios definidos para evaluar cada una de las

categorías, era evidente que la evaluación quedaba sujeta al punto de vista de cada especialista.

Se advirtió además que, si bien el *Manual* ofrece lineamientos claros con respecto a cada sección, la escala de 1 al 5 representaba una escala de claroscuros, que no tenía asidero con los lineamientos del *Manual*. Así pues, se determinó que la escala de evaluación debía circunscribirse a dos opciones: “Sí cumple” o “No cumple”.

Tercera etapa

Después de esta modificación se realizó la Prueba piloto 2, que consistió en que las 3 expertas una vez más evaluaran de manera independiente los artículos D, E y F en lo concerniente a la categoría de *Aspectos generales*. La prueba dio como resultado un porcentaje de 90% de acuerdos contra 10% de desacuerdos. Con esto se concluyó que se habían identificado los parámetros y la escala adecuados para calificar dicha categoría.

Con relación al apartado de *Sistema de citación*, se determinó que para evaluar cada una de las categorías, las revisoras debían emplear las siguientes tablas y valorar cada caso en base en la opción “Sí cumple” o “No cumple” (Tablas 3, 4 y 5).

Tabla 3. Matriz de evaluación de las citas y referencias

Caso	Sí cumple	No cumple
1.		
2.		
3.		
4.		
5.		
6.		
7.		
8.		
9.		
10.		
11.		

Tabla 4. Matriz para las referencias en el texto

Caso	Sí cumple	No cumple
1.		
2.		
3.		
4.		
5.		
6.		
7.		
8.		
9.		
10.		
11.		

Tabla 5. Matriz para la lista de referencia

Caso	Sí cumple	No cumple
1.		
2.		
3.		
4.		
5.		
6.		
7.		
8.		
9.		
10.		
11.		

Las indicaciones señalaban que debían evaluarse todas y cada una de las citas y paráfrasis del artículo, todas las referencias en el texto y la lista completa de referencias. Así, las gráficas anteriores se podrían ampliar según los casos registrados. Se proponía que cada caso se numerara en el artículo y que se evaluara en la tabla correspondiente. Para obtener el resultado final en una escala de 10/0 puntos se realizaría una regla de tres para cada una de las tres categorías evaluadas: citación directa y paráfrasis, referencias en el texto y lista de referencias.

Cuarta etapa

En base a los parámetros anteriores, se llevó a cabo la Prueba piloto 3, la cual consistió en evaluar el artículo G. Los resultados arrojaron un 25% de acuerdos contra un 75% de desacuerdos. Al analizar los resultados, se encontró que se estaban aplicando criterios diferentes para definir “Sí cumple” o “No cumple”. Por ello, se definieron parámetros específicos para la citación directa y la paráfrasis, así como para las referencias en el texto y para la lista de referencias (ver Tablas 3 y 5 del Anexo).

La última dificultad identificada en esta etapa fue que la labor de revisar la totalidad de las citas, paráfrasis, referencias y la lista de referencias del artículo requería muchas horas de trabajo. Como el objetivo era diseñar un instrumento ágil, se determinó que solo se revisarían 10 casos de citas directas o paráfrasis, 10 casos de referencias en el texto y 10 referencias tomadas de la lista de referencias. A fin de que las tres expertas evaluaran el mismo objeto, se decidió que los casos seleccionados siempre fueran los primeros 10 identificados a partir de la introducción (para *Citación directa y paráfrasis*, así como *Referencias en el texto*) y los primeros 10 casos de la *Lista de referencias*. Se estableció que si el artículo no contaba con los 10 casos en alguna de las categorías, se realizaría una regla de tres en función de los casos identificados. También se determinó que si el artículo no contaba con ningún caso (como suele ocurrir en la categoría de *Citación directa y paráfrasis*), al término de la tabla se consignaría la leyenda “Sin evaluar”.

Quinta etapa

Al llegar a este punto, se llevó a cabo la Prueba piloto 4, que consistió en que las tres revisoras evaluaran los artículos H, E e I. Los resultados arrojaron un 90% de acuerdos contra un 10% de desacuerdos, sin que surgieran discrepancias significativas. Con esto se concluyó que se habían identificado los criterios y la escala adecuados para calificar la categoría de Sistema de citación.

Sexta etapa

Finalmente, se determinó realizar una Prueba piloto 5, uniendo las secciones de *Aspectos generales* y *Sistema de citación*. La prueba consistió en que las 3 expertas evaluaran el artículo J. Los resultados arrojaron un 90% de acuerdos contra un 10% de desacuerdos. Entonces se procedió a verificar la fiabilidad y validez del instrumento.

Al confrontar los resultados, se detectó la necesidad de asegurar en las tres categorías del *Sistema de citación* (citación y paráfrasis, referencias en el texto y lista de referencias) que se estaba aludiendo al mismo caso y aplicando el mismo criterio para evaluarlo. Por ello, se determinó añadir en cada tabla una sección denominada *Nombre del autor o autores* a fin de ubicar en el artículo el caso en cuestión y una sección de “Comentarios”, donde se especificaran las razones por las que no se había cumplido un determinado criterio. A continuación se muestran algunos ejemplos:

Tabla 6. Referencias en el texto

Caso	Nombre del autor o autores	Sí cumple	No cumple	Comentarios
1.	PISA		X	No se escribió el nombre completo de la institución antes de la abreviatura, como lo determina el <i>Manual</i> cuando se trata de una primera mención. (APA, 2010, p. 177).
2.	(Aravena y Caamaño, 2013)		X	Por tratarse de autores que están entre paréntesis, se debió haber empleado el <i>ampersand</i> en vez de la y.

Tabla 7. Lista de referencias

Caso	Nombre del autor o autores	Sí cumple	No cumple	Comentarios
1.	Afonso, M. C.		X	Como se trata de una tesis se debió haber colocado primero el nombre de la institución y después el lugar (APA, 2010, p. 207).
2.	Aires, A. P., Campos, H. y Poças, R.		X	En lugar de la y, se debió haber empleado la coma y el <i>ampersand</i> antes del último autor. Además faltó cursiva en el número de volumen y no debió haberse añadido la abreviatura “pp” (<i>Manual</i> , p. 186).

Después de apreciar las bondades de estas inclusiones, se decidió que sería muy útil conservarlas en la rúbrica a fin de dar a conocer al autor, de manera puntual, la ubicación precisa del caso y la razón por la que no había logrado el cumplimiento del criterio.

Habiendo llegado a este punto del proceso, se consideró imprescindible determinar la validez y fiabilidad de la rúbrica. El primer paso consistió en evaluar la validez de contenido del instrumento, lo cual se logró gracias al amplio conocimiento que las expertas tienen sobre el uso de los lineamientos APA. Posteriormente, y tal como se ha mostrado en los apartados anteriores, se efectuó la validación de las dimensiones del instrumento para determinar la validez teórica de los constructos. Este riguroso proceso se basó no solo en un estricto apego al *Manual*, sino también en el permanente intercambio de puntos de vista entre las expertas.

Una vez efectuada la validez cualitativa del instrumento, se procedió a determinar su fiabilidad. De acuerdo con Cooksey (2014) y Gwet (2014) es crucial determinar la fiabilidad de un instrumento ya que diferentes revisores pueden interpretar el mismo dato de manera distinta. En consecuencia, con el fin de verificar estadísticamente los resultados obtenidos, se realizó la prueba Fleiss’ kappa, la cual mide la proporción de acuerdo entre los observadores. Fleiss’ kappa es una extensión de Cohen kappa, que es utilizada cuando el número de observadores es mayor a dos y existe un número fijo de elementos. En esta prueba los coeficientes de correlación varían entre 0 y 1, donde 0 significa que no hay acuerdo y 1 representa acuerdo perfecto. Se considera que un valor de Fleiss’ kappa igual o superior a .7 representa que el instrumento es fiable. Al realizar la prueba se obtuvo un valor de $k=.7$, lo que indica que la fuerza del acuerdo

entre las revisoras fue consistente y por tanto el instrumento resulta fiable.

Tabla 8. Fleiss' Kappa, proporción de acuerdo entre los expertos

Proporción de acuerdo de los elementos	Proporción de acuerdo a las categorías	P	P e	K
0.25	1.00	.72	0.06	.7
-	0.17			
-	1.00			
0.42				

Una vez que se concluyó que la rúbrica era fiable y válida, se redactaron las instrucciones para su manejo. La rúbrica se presenta de manera independiente en el Anexo 1 a fin de que quede a disposición de los interesados.

4 DISCUSIÓN Y CONCLUSIONES

Esta investigación se propuso inicialmente diseñar y validar una rúbrica para evaluar artículos científicos conforme a su adecuación al estilo APA. Durante el proceso se hicieron varios hallazgos, que dieron como consecuencia la configuración final de la rúbrica y el replanteamiento del objetivo inicial.

El primer punto del proceso fue reconocer la imposibilidad de elaborar una rúbrica que evaluara todos los elementos de esta normativa. Por ello fue necesario aceptar que el instrumento no evaluaría todos los aspectos del estilo APA, sino solo algunos. El segundo punto fue identificar las categorías susceptibles de evaluación y los parámetros en base a los cuales se llevaría a cabo dicha evaluación. Una vez confirmadas las categorías y los parámetros, y después de las pruebas piloto, se llegó a la conclusión de que no era viable revisar todos los elementos del artículo correspondientes al sistema de citación, ya que esta labor requeriría muchas horas de trabajo para el revisor. Por ello se determinó que para aportar una rúbrica ágil era necesario reducir el tamaño de la muestra a 10 casos por categoría.

Después de efectuar las pruebas necesarias, se llegó a la conclusión de que se disponía de un instrumento fiable que, sin embargo, presentaba algunas limitaciones: el instrumento no era capaz de evaluar todos los aspectos del estilo APA contenidos en un artículo científico, ni de evaluar el sistema de citación de un artículo en su totalidad. No obstante, a pesar de las limitaciones antes enumeradas, se concluyó que la rúbrica tenía el mérito de poner ante los ojos del revisor y, de manera gráfica y puntual, criterios clave que a veces pueden perderse de vista en el vasto océano del Manual. Del mismo modo, la rúbrica podría serle útil al autor para identificar sus fortalezas y debilidades en cuanto a su manejo del estilo APA.

Según la experiencia profesional de las expertas, se manejan diferentes criterios en la aplicación del estilo APA, incluso entre investigadores de mismo centro de investigación o académicos de una misma universidad. Esta diversidad de criterios provoca pérdidas de tiempo y un gran desgaste: polémicas entre colegas, conflictos entre autores y revisores o correctores de estilo. Así, se espera que al poner sobre la mesa algunos elementos clave del estilo APA la presente rúbrica contribuya a incrementar la coherencia y uniformidad en el proceso de redacción y revisión de los artículos científicos, y que sea un apoyo al trabajo de autores, editores y evaluadores de artículos científicos.

Uno de los hallazgos del proceso fue advertir que la rúbrica había adquirido una función diferente, adicional a la prevista en un inicio. Más que arrojar una calificación sin más, el

instrumento desveló tener una función didáctica ya que muestra los aspectos clave que han de tenerse en cuenta y retroalimenta a los autores sobre los errores que deben corregir.

Al emplear la rúbrica y comparar los resultados, también se hizo evidente que podía ser muy útil para que los revisores, compartiendo sus resultados, perfeccionaran su conocimiento del estilo APA. Por otra parte, se llegó a la conclusión de que el instrumento podría contribuir a la unificación de criterios en instituciones académicas. Así pues, la rúbrica podría emplearse con éxito para la enseñanza del estilo APA en cursos de formación a investigadores y profesores, y los profesores, por su parte, podrían emplear la rúbrica para realizar prácticas con sus alumnos en la enseñanza del estilo APA.

Otra de las bondades que ofrece la rúbrica es la de contar con la flexibilidad para ajustarse a las necesidades de cada usuario. Por ejemplo, en el apartado del Sistema de citación la muestra podría ampliarse a dos o más ensayos. En el caso de los maestros, ellos podrían modificar el apartado de *Aspectos generales* con los parámetros que consideraran pertinentes para evaluar los ensayos académicos de sus alumnos.

Finalmente, se espera que la presente investigación dé pie a estudios posteriores por medio de los cuales el instrumento actual sea enriquecido y aplicado en diferentes ámbitos de la escritura científica o académica en sus diferentes géneros.

REFERENCIAS

- American Psychological Association (2010). *Manual de Publicaciones de la American Psychological Association* (3ª ed.). México: Manual Moderno.
- Andrade, H. G. (2005). Teaching with rubrics: The good, the bad, and the ugly. *College Teaching*, 53(1), 27–30. doi:10.3200/CTCH.53.1.27-31
- Brown, G. T. L., Glasswell, K., & Harland, D. (2004). Accuracy in the scoring of writing: Studies of reliability and validity using a New Zealand writing assessment system. *Assessing Writing*, 9(2), 105–121. doi:10.1016/j.asw.2004.07.001
- Burt, C. G., Cima, R. R., Koltun, W. A., Littlejohn, C. E., Ricciardi, R., Temple, L. K., Rothemberger, D. A., & Baxter, N. N. (2009). Developing a research agenda for the American Society of Colon and Rectal Surgeons: Results of a Delphi approach. *Diseases of the Colon Rectum*, 52, 898–905. doi:10.1007/DCR.0b013e3181a0b358
- Camacho, M. E., Rojas, M. E., & Rojas, L. (2014). El artículo científico para revista académica: pautas para su planificación y edición de acuerdo con el modelo APA. *e-Ciencias de la Información*, 4(2), 1–29.
- Cho, K., Schunn, C. D., & Wilson, R. W. (2006). Validity and reliability of scaffolded peer assessment of writing from instructor and student perspectives. *Journal of Educational Psychology*, 98(4), 891–901. doi:10.1037/0022-0663.98.4.891
- Cooksey, R. (2014). *Illustrating Statistical Procedures. Finding Meaning in Quantitative Data*. University of New England, Australia: Tilda University Press.
- De Villiers, M. R., De Villiers, P. J., & Kent, A. P. (2005). The Delphi technique in health sciences education research. *Medical Teacher*, 27(7), 639–643. doi:10.1080/13611260500069947
- Floyd, R. G., Cooley, K. M., Arnett, J. E., Fagan, T. K., Mercer, S. H., & Hingle, C. (2011). An overview and analysis of journal operations, journal publication patterns, and journal impact in school psychology and related fields. *Journal of School Psychology*, 49(6), 617–647. doi:10.1016/j.jsp.2011.11.008
- Giménez, E. (2014). Imposturas en el ecosistema de la publicación científica. *Revista de Investigación Educativa*, 32(1), 13–23. doi:10.6018/rie.32.1.190251
- Greenberg, K. P. (2012). A reliable and valid weighted scoring instrument for use in grading APA-Style empirical research report. *Teaching of Psychology*, 39(1), 17–23. doi:10.1177/0098628311430643
- Gwet, K. L. (2014). *Handbook of inter-rater reliability: The definitive guide to measuring the extent of agreement among raters*. Advanced Analytics, LLC.
- Hafner, J. C., & Hafner, P. M. (2003). Quantitative analysis of the rubric as an assessment tool: An empirical study of student peer-group rating. *International Journal of Science Education*, 25, 1509–1528. doi:10.1080/0950069022000038268
- Hernández, M., Fernández, C., & Baptista, P. (2003). *Metodología de la investigación* (3ª ed.). México: McGraw-Hill.

- Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review*, 2, 130–144. doi:10.1016/j.edurev.2007.05.002
- Moskal, B. M. (2000). Scoring rubrics: What, when and how? *Practical Assessment, Research & Evaluation*, 7(3), 1–5.
- Moskal, B. M., & Leydens, J. A. (2000). Scoring rubric development: Validity and reliability. *Practical assessment, research & evaluation*, 7(10), 71–81.
- Newell, J. A., Dahm, K. D., & Newell, H. L. (2002). Rubric development and inter-rater reliability issues in assessing learning outcomes. *Chemical Engineering Education*, 36(3), 212–215.
- Nickerson, R. S. (2005). What authors want from journal reviewers and editors. *American Psychologist*, 60, 661–662. doi:10.1037/0003-066X.60.6.661
- Peat, B. (2006). Integrating writing and research skills: Development and testing of a rubric to measure student outcomes. *Journal of Public Affairs Education*, 12, 295–311.
- Picón, E. (2013). The Role of Rubrics in Fair Assessment Practices. *Íkala, Revista de Lenguaje y Cultura*, 18(3), 7–994.
- Roddy, E., Zhang, W., Doherty, M., Arden, N. K., Barlow, J., Birrell, F., Carr, A., Chakravarty, K., Dickson, J., Hay, E., Hosie, G., Hurley, M., Jordan, K. M.; McCarthy, C., McMurdo, M., Mockett, S., O'Reilly, S., Peat, G., Pendleton, A., & Richards, S. (2006). EvidenceBased clinical guidelines: a new system to better determine true strength of recommendation. *Journal of Evaluation in Clinical Practice*, 12(3), 347–352. doi:10.1111/j.1365-2753.2006.00629.x
- Stellmack, M. A., Konheim-Kalkstein, Y. L., Manor, J. E., Massey, A. R., & Schmitz, J. A. P. (2009). An assessment of reliability and validity of a rubric for grading APA-style introductions. *Teaching of Psychology*, 36(2), 102–107. doi:10.1080/00986280902739776
- Steurer J. (2011). The Delphi method: an efficient procedure to generate knowledge. *Skeletal Radiol*, 40(8), 959–61. doi:10.1007/s00256-011-1145-z
- Thaler, N., Kazemi, E., & Huscher, C. (2009). Developing a rubric to assess student learning outcomes using a class assignment. *Teaching of Psychology*, 36(2), 113–116. doi:10.1080/00986280902739305

AGRADECIMIENTOS

Esta investigación se ha desarrollado en el marco del Grupo de Investigación Interdisciplinar en Docencia Universitaria (GIDU) y del Proyecto Diseño y Atención a las Oportunidades de Género en la Educación Superior de la Universidad de Alicante.

Con el fin de llegar a un mayor número de lectores, NAER ofrece traducciones al español de sus artículos originales en inglés. **Este artículo en español no es la versión original del mismo, sino únicamente su traducción.** Si quiere citar este artículo, por favor, consulte el artículo original en inglés y utilice la paginación del mismo en sus citas. Gracias.

ANEXO 1

RÚBRICA

Esta rúbrica es un instrumento que se ha diseñado con el objetivo de ofrecerle al autor o autores de artículos científicos una primera evaluación con respecto a su manejo del estilo APA conforme a las pautas del *Manual de Publicaciones de la American Psychological Association* (en lo sucesivo el *Manual*), en su 3ª edición en español, publicada en 2010.

Se comprende que la rúbrica no ofrece una evaluación total del artículo, ni una evaluación completa de los aspectos revisados. Concretamente la rúbrica considera dos apartados que hemos denominado: *Aspectos generales* y *Sistema de citación*. La siguiente tabla ilustra las categorías a evaluar en cada uno de estos apartados:

Tabla 1. Categorías de los Aspectos generales y del Sistema de citación

Aspectos generales	Sistema de citación
Título	Citación directa y paráfrasis
Secciones	Referencias en el texto

Nombre del autor	afiliación	Lista de referencias
Resumen		
Introducción		
Método		
Resultados		
Comentarios (Discusión)		
Estructura del artículo		
Lenguaje empleado	libre de	discriminaciones

A continuación se proporcionan las instrucciones para cumplimentar las tablas que permiten evaluar dichas categorías. A lo largo de la evaluación es imprescindible la consulta del *Manual*. Entre paréntesis se indicará el número de página o páginas de referencia.

INSTRUCCIONES

Aspectos generales

Los aspectos generales se revisarán conforme a 10 parámetros que se describen en la siguiente tabla.

Cada uno de los parámetros obtendrá un puntaje de 1 si cumple y de 0 si no cumple. En el apartado correspondiente el revisor añadirá los comentarios que considere pertinentes para retroalimentar al autor. Al final de la tabla se registrará el total de puntos obtenidos.

Tabla 2. Tabla para evaluar los Aspectos generales

Aspectos generales	Sí cumple	No cumple	Comentarios
1. El título sintetiza la idea principal del escrito de una manera sencilla y con un estilo adecuado.			
2. El artículo contiene las secciones propuestas por el estilo APA: resumen, introducción, método, resultados, comentarios (discusión) ¹ .			
3. El nombre de autor y la afiliación institucional se ciñe a los criterios indicados por el estilo APA.			
4. El resumen es preciso, no evaluativo, coherente, legible y conciso.			
5. La introducción plantea el problema y describe la estrategia de la investigación.			
6. La sección de método incluye una descripción completa de los métodos empleados.			
7. La sección de resultados presenta los datos relevantes que sustentarán las conclusiones.			
8. La sección de comentarios (discusión) examina, interpreta y califica los resultados, enfatizando sus consecuencias teóricas o prácticas.			
9. El artículo tiene una estructura, caracterizada por la utilización de diferentes niveles de encabezados para indicar la jerarquización de la información.			

10. El lenguaje empleado está libre de discriminaciones a grupos o a personas.			
Total:			

Nota: En el caso de que alguna sección no se incluya porque el artículo, debido a su naturaleza, no tiene que presentarla, se categorizará como “Sí cumple”ⁱ.

Citación directa y paráfrasis

El revisor analizará 10 casos (ya sean citas directas menores a 40 palabras, citas en bloque o paráfrasis), los cuales serán los primeros localizados en el artículo a partir de la introducción.

A continuación se describen los parámetros para evaluar esta categoría:

Tabla 3. Parámetros para evaluar la *Citación directa y la paráfrasis*

Sí cumple	
<ul style="list-style-type: none"> Las citas directas (menores a 40 palabras) deben contener los siguientes elementos indicados por el estilo APA: autor, año, página específica de la cita y las comillas de apertura y de cierre. Independientemente de los distintos formatos, el número de página debe ubicarse al final de la citaⁱⁱ, precedido de una coma. Las citas mayores de 40 palabras deben disponerse en bloque y sin comillas (APA, 2010, p. 171). Al término del bloque debe consignarse el autor y la página específica de la cita o únicamente la página cuando el autor fue mencionado antes. Para efectos de esta rúbrica se considerarán “paráfrasis” aquellas que contengan no solo autor y año, sino también paginaciónⁱⁱⁱ. Los casos que sigan el formato “autor-fecha” se revisarán en la categoría <i>Referencias en el texto</i>. Cuando se citen entre paréntesis dos autores, se deberá emplear el símbolo del <i>ampersand</i> (&). Cuando se citen entre paréntesis más de tres autores, se deberá emplear la coma antes del símbolo del <i>ampersand</i> (&). 	
Nota:	
<ul style="list-style-type: none"> El uso de mayúsculas o minúsculas en la citación directa o en la paráfrasis no será objeto de evaluación. 	
No cumple	
<ul style="list-style-type: none"> Las citas directas carecen de algunos de los siguientes elementos: autor, año, número de páginas o el empleo de comillas. Las citas mayores de 40 palabras no están dispuestas en bloque o presentan un bloque entre comillas. Al término del bloque no se consigna la página específica de la cita. Las paráfrasis carecen de autor o año. Falta del <i>ampersand</i> entre autores que se citan entre paréntesis. 	

Cada uno de los casos se evaluará conforme a los criterios “Sí cumple” o “No cumple”. A fin de que el autor del artículo pueda ubicar el caso en cuestión, el revisor consignará el nombre del autor o autores del caso revisado. En la sección de “Comentarios” se especificarán las razones por las que en un momento dado no se ha cumplido con un determinado parámetro.

A continuación se presenta la tabla para evaluar esta categoría:

Tabla 4. Tabla para evaluar la *Citación directa y la paráfrasis*

Caso	Nombre del autor o autores	Sí cumple	No cumple	Comentarios
1.				
2.				
3.				
4.				
5.				
6.				
7.				
8.				
9.				
10.				
Total:				

Si el artículo tiene un número menor a 10 casos, la calificación se obtendrá mediante una regla de tres. Si el artículo no contiene citas directas, citas en bloque o paráfrasis, este apartado no será sujeto a evaluación. En ese caso en el apartado de “Total” se consignará la leyenda “Sin evaluar”.

Referencias en el texto

El revisor analizará 10 casos, los cuales serán los primeros localizados en el artículo a partir de la introducción.

A continuación se describen los parámetros para evaluar esta categoría:

Tabla 5. Parámetros para evaluar las *Referencias en el texto*

Sí cumple	
Las referencias en el texto deben cumplir con el sistema de citas de autor-fecha, estipulado por el estilo de la APA (APA, 2010, pp. 174-179).	
A continuación se detallan algunos puntos que el revisor no debe perder de vista:	
<ul style="list-style-type: none"> En el cuadro de la página 177, el <i>Manual</i> indica puntualmente el formato a seguir para las referencias en el texto. Estas dependen de dos factores siempre correlacionados: a) el tipo de cita y b) si se trata de la primera referencia en el texto o si se trata de referencias subsecuentes (APA, 2010, p. 177). El tipo de cita alude a si se trata del trabajo de un solo autor, un trabajo de dos autores, etc. Con respecto al segundo factor habrá que considerar, además, si las referencias están o no entre paréntesis. Cuando se citen entre paréntesis dos autores o más autores se deberá emplear el símbolo del <i>ampersand</i> (&) antes del último autor. Cuando se citen entre paréntesis más de tres autores se deberá emplear la coma y el símbolo del <i>ampersand</i> (&) antes del último autor. En el caso de referencias entre paréntesis deberá insertarse una coma antes de la fecha. En caso de que se citen dos o más trabajos dentro del mismo paréntesis, se deben disponer alfabéticamente, conforme al apellido de los autores. En estos casos debe emplearse punto y coma entre los autores, como en el siguiente ejemplo: Diversos estudios (Allen, 2012; Miller, 1999; Safranske, 1998) afirman que... 	
No cumple	
<ul style="list-style-type: none"> Las referencias carecen de alguno de los siguientes elementos: autor o fecha. Las citas contravienen el formato indicado para cada caso. Se omite el empleo indicado del <i>ampersand</i> o de la coma para cada caso. 	

Cada uno de los casos se evaluará conforme a los criterios “Sí cumple” o “No cumple”. Se contará como un caso no solo las referencias en las que señale una obra (escrita por uno o varios autores), sino también aquellas referencias en las que se mencionen varios trabajos de distintos autores. Para una mejor comprensión se ofrecen los siguientes ejemplos:

Tabla 6. Ejemplos de referencias

Referencia 1	Kessler (2003)	Una sola obra de un solo autor
Referencia 2	Kisangau, Lyaruu, Josea y Joseph (2007)	Una sola obra de varios autores
Referencia 3	Diversos estudios (Allen, 2012; Miller, 1999; Safranske, 1998)	Varias obras de varios autores afirman que...

A fin de que el autor del artículo pueda ubicar el caso en cuestión, el revisor consignará el nombre del autor o autores del caso revisado. En la sección de “Comentarios” se especificarán las razones por las que en un momento dado no se ha cumplido con un determinado parámetro.

A continuación se presenta la tabla para evaluar esta categoría:

Tabla 7. Tabla para evaluar las *Referencias en el texto*

Caso	Nombre del autor o autores	Sí cumple	No cumple	Comentarios
1.				
2.				
3.				
4.				
5.				
6.				
7.				
8.				
9.				
10.				
Total:				

Lista de referencias

El revisor analizará 10 casos, los cuales serán los primeros localizados en la lista de referencias.

A continuación se describen los parámetros para evaluar esta categoría:

Tabla 8. Parámetros para evaluar la *Lista de referencias*

Sí cumple	
•	Cada referencia debe contar con los elementos indicados por el estilo de la APA y seguir el formato estipulado según el tipo de publicación (libro, artículo impreso, artículo en línea, tesis, etc. (APA, 2010, pp. 184-215).
•	Se deberán considerar que se cumpla el orden de las referencias en la lista.
•	Cuando se trata de dos autores o más, deberá colocarse una coma y el <i>ampersand</i> (&) antes del último autor.
Notas:	
•	El empleo de un espaciado o de una sangría diferente de lo dispuesto por el estilo de la APA en la lista de referencias no será objeto de evaluación.
•	El empleo de abreviaturas diferentes de lo dispuesto por el estilo de la APA no será objeto de evaluación.
No cumple	

- La referencia carece de alguno de los elementos indicados por el estilo de la APA o no sigue el formato estipulado según el tipo de publicación.

Cada uno de los casos se evaluará conforme a los criterios “Sí cumple” o “No cumple”. A fin de que el autor del artículo pueda ubicar el caso en cuestión, el revisor consignará el nombre del autor o autores del caso revisado. En la sección de “Comentarios” se especificarán las razones por las que en un momento dado no se ha cumplido con un determinado parámetro.

A continuación se presenta la tabla para evaluar esta categoría:

Tabla 9. Tabla para evaluar la *Lista de referencias*

Caso	Nombre del autor o autores	Sí cumple	No cumple	Comentarios
1.				
2.				
3.				
4.				
5.				
6.				
7.				
8.				
9.				
10.				
Total:				

Rúbrica

La siguiente tabla ilustra de manera global los diferentes niveles de dominio que puede alcanzar el autor en cada uno de los aspectos evaluados. El revisor sombreará sobre la tabla el nivel de puntaje obtenido en cada uno de los aspectos revisados y además el nivel de logro alcanzado (Muy bien, Bien, Regular, Deficiente) de acuerdo con la suma de los cuatro aspectos.

Tabla 10. Resultado global de la evaluación

	Aspectos generales	Citas directas y paráfrasis	Referencias en el texto	Lista de referencias
Muy bien (40-36)	(10 puntos) Se registra un muy buen manejo de los aspectos generales.	(10 puntos) Las 10 citas o entre el 99% y el 90% de ellas siguen los lineamientos de la APA.	(10 puntos) Las 10 referencias se consignan en el texto conforme a los estándares indicados por el estilo APA.	(10 puntos) Las 10 referencias siguen los estándares indicados por el estilo APA.
Bien (35-31)	(9-8 puntos) Se registra un buen manejo de los aspectos generales.	(9-8 puntos) Entre 9 y 8 citas o entre el 89% y el 80% de ellas siguen los lineamientos de la APA.	(9-8 puntos) Entre 9 y 8 de las referencias se consignan en el texto conforme a los estándares indicados por el estilo APA.	(9-8 puntos) Entre 9 y 8 de las referencias siguen los estándares indicados por el estilo APA.
Regular (30-26)	(7-6 puntos) Se registra un	(7 a 6 puntos) Entre 7 y 6 citas o entre	(7 a 6 puntos) Entre 7 y 6 de las refe-	(7 a 6 puntos) Entre 7 y 6 de las

	manejo regular de los aspectos generales.	el 79% y el 70% de las citas siguen los lineamientos APA.	referencias se consignan en el texto conforme a los estándares indicados por el estilo APA.	referencias siguen los estándares indicados por el estilo APA.
Deficiente (menos de 25)	(5 puntos o menos) Se registra un manejo deficiente de los aspectos generales.	(5 puntos o menos) Cinco citas o menos del 69% de ellas siguen los lineamientos APA.	(5 puntos o menos) Menos de 6 referencias se consignan en el texto conforme a los estándares indicados por el estilo APA.	(5 puntos o menos) Menos de 6 de las referencias siguen los estándares indicados por el estilo APA.

Se comprende que en caso de que uno de los aspectos no haya sido evaluado por falta de casos, el puntaje se ajustará para obtener el nivel de logro alcanzado.

NOTAS

ⁱ El *Manual* no estipula que los artículos científicos deban contar estrictamente con dichos apartados. De hecho, afirma: “El *Manual de publicaciones* presenta requerimientos de estilo explícitos, pero reconoce que las alternativas a veces son necesarias. Los autores deben equilibrar las reglas del *Manual*.”

ⁱⁱ Aunque el *Manual* no lo afirma de manera explícita, queda claro en todas sus ejemplificaciones que el número de página se ubica al final de cada citación (ya sea cita directa o paráfrasis). En <http://blog.apastyle.org/apastyle/2010/03/how-to-cite-direct-quotations.html> se ofrecen distintos formatos de citación que confirman este criterio.

ⁱⁱⁱ El *Manual* indica que en el caso de la paráfrasis “se aconseja indicar un número de página o párrafo” (APA, 2010, p. 170), pero se comprende que este dato no es obligatorio.